Markov dynamics as a zooming lens for multiscale community detection: non clique-like communities and the field-of-view limit

Michael T Schaub[1,2,*], Jean-Charles Delvenne[3], Sophia N Yaliraki[2] and Mauricio Barahona[1,*]

[1] Department of Mathematics, Imperial College London, London, United Kingdom
[2] Department of Chemistry, Imperial College London, London, United Kingdom
[3] Department of Applied Mathematics, Université catholique de Louvain, Louvain-la-Neuve, Belgium
[*] E-mail: michael.schaub09@imperial.ac.uk, m.barahona@imperial.ac.uk

Abstract

In recent years, there has been a surge of interest in community detection algorithms for complex networks. A variety of computational heuristics, some with a long history, have been proposed for the identification of communities or, alternatively, of good graph partitions. In most cases, the algorithms maximize a particular objective function, thereby finding the 'right' split into communities. Although a thorough comparison of algorithms is still lacking, there has been an effort to design benchmarks, i.e., random graph models with known community structure against which algorithms can be evaluated. However, popular community detection methods and benchmarks normally assume an implicit notion of community based on clique-like subgraphs, a form of community structure that is not always characteristic of real networks. Specifically, networks that emerge from geometric constraints can have natural non clique-like substructures with large effective diameters, which can be interpreted as long-range communities.

In this work, we show that long-range communities escape detection by popular methods, which are blinded by a restricted 'field-of-view' limit, an intrinsic upper scale on the communities they can detect. The field-of-view limit means that long-range communities tend to be overpartitioned. We show how by adopting a dynamical perspective towards community detection [1, 2], in which the evolution of a Markov process on the graph is used as a zooming lens over the structure of the network at all scales, one can detect both clique- or non clique-like communities without imposing an upper scale to the detection. Consequently, the performance of algorithms on inherently low-diameter, clique-like benchmarks may not always be indicative of equally good results in real networks with local, sparser connectivity. We illustrate our ideas with constructive examples and through the analysis of real-world networks from imaging, protein structures and the power grid, where a multiscale structure of non clique-like communities is revealed.

Introduction

The analysis of community structure in complex networks has gained much attention in recent years and a variety of community detection algorithms have been proposed (for a recent overview see Ref. [3]). The reason for this interest is that by finding community structure in large networks one hopes to reveal relevant modules at mesoscopic scales that can affect or explain the global behavior of the system. Community detection may thus facilitate new insights into the structural and functional organization of a system (and the interplay between these two), as well as potentially serving as the basis for reduced descriptions of complex systems. Examples of important applications include networks from technological, physico-chemical, biological and medical data as well as data from the social sciences (see e.g. [3–6] and references therein).

Community detection algorithms are based on diverse notions of what makes a good community. Different mathematical and computational heuristics, some of them based on graph partitioning concepts, have been used to identify communities or to obtain an optimized split of the original network into

smaller subgraphs with a community-like character. A common trait in many algorithms is to group nodes based on edge density: communities concentrate high edge weight within them, while having low edge weight between them. This structural notion has led to several heuristics including, among others, modularity [7,8] and multiscale-Potts models [9–11] which have been coupled with different optimization algorithms (e.g., the Louvain method [12]) for the maximization of the corresponding cost functions. Recently, the Map equation framework has proposed an alternative notion that views communities as groupings of nodes that lead to concise descriptions (in an information-theoretical sense) for the process of communicating the position of a random walker within the network [13,14].

To aid in the comparative evaluation of community detection algorithms, there has also been an effort to design benchmark graph models with an embedded community structure [7, 15–17]. However, in designing these benchmarks, a particular notion of community has to be adopted implicitly. As in many of the community detection algorithms described above, such benchmark models are based on the customary structural notion of community in terms of edge density. Therefore, the community structure is introduced as a *stochastic clique-of-cliques*, i.e., as a hierarchy of interconnected realizations of Erdös-Renyi graphs with block-wise homogeneous, 'all-to-all' edge densities. This notion is motivated by the fact that many complex systems in the literature have been found to display small diameters and is indeed applicable to a wide range of networks, e.g. networks constructed from correlations or networks from the social sciences.

However, as we show below, there is a broad class of locally sparser networks where the assumption of 'all-to-all' connectivity is not warranted yet may still contain relevant modular substructures, which we would like to identify via community detection. A community in this case corresponds to a set of nodes which have a stronger direct or *indirect* connection with each other than with nodes outside their community. Such long-range substructures cannot be modelled accurately by stochastic cliques. This is the case in a variety of systems (e.g. biological and engineering networks) where entities are coupled in a complex manner via a chain of local interactions such that not all entities of a module are directly connected, yet they are more strongly related to each other than to entities from a different subsystem. Therefore, the evaluation of algorithms on benchmarks with clique-like communities might not be representative of their ability to identify relevant structures on real networks with localized, sparser connectivity where non-clique like communities may still exist.

The characteristics of communities in such networks (relatively sparse with a strong local structure) also highlight a limitation in popular community detection methods. It is already known that some of these methods (e.g., modularity) are affected by a 'resolution limit' [18], a lower scale that establishes a minimum size below which communities cannot be detected. Infomap, on the other hand, seems to be immune to such a limit [19]. We show below that these methods also have a *'field-of-view' limit*, an upper limit in the effective diameter of the communities they can detect. This field-of-view limit affects both modularity and Infomap. Therefore, such structural methods contain an implicit scale and can only detect communities that lie within a range of effective sizes (dependent on each method and each graph) and might miss groups outside of it.

One way to correct for the field-of-view limit is to adopt an approach to community detection that intrinsically *scans across scales*, such as the partition stability framework [1, 2]. Stability is a recently proposed dynamical approach to community detection which follows the time evolution of a Markov diffusion process on the graph incorporating increasingly longer paths on the network. This process entails a natural dynamic sweeping that can be understood as the application of a zooming lens to community detection. It is important to remark that the sweeping process is inherent and key to the methodology: the dynamical basis implies a systematic tool by which the structure at all scales needs to be examined, without looking for the 'right' scale. In doing so, stability is able to reveal communities without imposing a scale a priori or, indeed, community structure at multiple scales. It has been shown [1] that in such a framework, the standard structural notion of community given by modularity is a particular case based on one-step transitions (i.e., detected at Markov time equal to 1). It can be shown [20] that

Infomap also considers one-step transitions in an averaged manner.

Under this interpretation, it is easy to understand why one-step methods (such as modularity and Infomap) cannot detect communities with large effective diameter, since these communities cannot be properly explored within one-step transitions. When faced with such communities, one-step algorithms tend to overpartition them. In fact, this feature provides us with an indicator that one-step methods are being applied to a community structure outside their range of applicability: when one-step methods return communities with large effective diameters, this can be seen as an indicator that the methods are operating on a graph which does not conform to their intrinsic notion of communities as cliques. On the other hand, long range communities can be revealed at longer Markov times through the dynamic sweeping provided by stability. We show below the relevance of these observations through the analysis of constructive examples and of real networks from biology, computer science and engineering.

Methods

Notation

Networks are here defined as undirected, connected, weighted graphs with N nodes. For simplicity, we consider non-bipartite graphs. The connectivity of the graph is encoded by the weighted adjacency matrix A, a symmetric matrix where $A_{ij} = A_{ji}$ corresponds to the weight of the edge between nodes i and j. The total weight of the edges is $m = \sum_{i \geq j=1}^{N} A_{ij}$.

Community detection methods: structural and dynamical interpretations

The purpose of this section is to provide a dynamical re-interpretation of some popular community detection algorithms that are based on structural notions of community. It will be shown that such methods can be seen as one-step methods. When reinterpreted as one-step dynamical methods, it becomes possible to understand the inherent assumptions or bias of structural methods towards the identification of short-range communities and the limits that this imposes on the detection of non clique-like communities.

A different interpretation of a well-known metric: Modularity as a one-step method

The well-known modularity [7,8,21] has in recent years been used as one of the standard metrics to evaluate and optimize community structure. The original idea of modularity was intrinsically combinatorial: it assigns nodes to communities such that the density of links inside a community is maximized, when compared to a random network with the same degree sequence. It has been shown previously [1] that this measure can be interpreted dynamically as a one-step method. Modularity, Q, can be rewritten as [1]:

$$Q = \text{trace } H^T \left[\Pi M - \pi^T \pi \right] H, \tag{1}$$

where $M = D^{-1}A$ is the one-step transition matrix of a discrete time random walk defined on the graph; D is a diagonal matrix with the strengths of the nodes on its diagonal ($D_{ii} = d_i = \sum_j A_{ij}$); $\pi_i = d_i/2m$ is the equilibrium distribution of the process; and $\Pi = \text{diag}(\pi)$. Furthermore the hard partitioning of the graph into c communities is encoded into a $N \times c$ indicator matrix H with $H_{ij} \in \{0, 1\}$, where a 1 denotes that node i belongs to community j.

Recasting modularity in this way suggests the following dynamical reinterpretation: communities are assigned such that the overall probability of starting in a community and remaining there after one step is maximized compared to the probability of randomly ending up in the same community when at equilibrium. Hence modularity can effectively be seen as a one-step method and, as such, it tends to favor communities in which the Markov process spreads rapidly, such as clique-like groups with weak connections between them. On the other hand, when the underlying communities are not clique-like (or

have a large effective intra-community distance), modularity optimization can lead to overpartitioning, i.e., to the detection of artificially small communities. This is shown in the examples in the Results section.

Another one-step method: Infomap and clique-like communities

Recently, Rosvall and Bergstrom [14] have proposed the (Info-)Map framework for community detection. The general idea behind the Map equation algorithm is to find a binary code with unique codewords for each node within a community that can be used to describe compactly the position of a random walker in the network. In a network with a marked community structure, the probability of flipping membership is smaller. Therefore, one can compress the description by reusing short node names (codewords) within separate communities, similarly to the way street names may be used in different cities throughout a country [14].

Although different in spirit to modularity, it is interesting to note that the Map equation, much like modularity, is also based solely on the equilibrium distribution π and a one-step process that reflects inter-community transitions in an block-averaged manner. Indeed, it can be shown that the Map equation formalism does not distinguish different connectivity structures inside communities [20]. This reflects an implicit assumption about fast mixing communities and, consequently, a homogeneous 'all-to-all' connection pattern (i.e., clique-like communities in the sense of having short effective diameter) matches best the inherent notion of community of the algorithm. Hence, Infomap is prone to finding communities that can be well approximated as stochastic cliques. Indeed, Infomap performed as one of the best algorithms in a recently conducted benchmark test where communities are defined as stochastic cliques [19]. This also explains why, unlike modularity, the Map equation is not known to be affected by the resolution limit: the fact Infomap is greedy with respect to locally dense substructures makes it well suited for detection even at the finest scales. However, this desirable, designed-for feature also means that when the graph under consideration has slowly mixing communities, Infomap displays a pronounced overpartitioning effect, even more so than modularity.

Stability as a dynamical framework for community detection: sweeping across scales

Communities with large effective diameters, e.g., non clique-like communities, can be missed by the methods above because they may not be discernible with one-step measures. This raises the question of how to detect non clique-like communities? Intuitively, a means to account for their structure is to consider walks of lengths greater than one. This can be done in a principled way using the recently proposed partition stability framework [1, 2]. The idea underpinning stability is to define a Markov (diffusion) process on the graph and follow how the probability flow spreads out over time. Rather than looking at one-step measures, stability takes into account walks of increasing length systematically by looking at larger times and, in doing so, it can reveal communities at different scales: in general, the longer the time, the coarser the partition. Alternatively, one can think of the Markov time as providing a *zooming lens* that scans the structure of the graph from the finer to the coarser structure. Importantly, stability has been shown to provide a unifying framework for several measures of community detection including modularity and spectral partitioning [1]. In the following, a brief introduction to the stability measure will be given. For a more detailed exposition see [1, 2].

The stability of the partitioning of a graph can be defined via the autocovariance of a Markov process taking place on the graph. The definition can be based on both discrete and continuous times. Here we consider the continuous-time Markov process on the graph governed by the following (Laplacian) dynamics:

$$\dot{\mathbf{p}} = -\mathbf{p}\,[D^{-1}L], \tag{2}$$

where $\mathbf{p}$ is a $1 \times N$ probability vector, D is the diagonal matrix with the strengths d_i, and $L = D - A$ is the graph (combinatorial) Laplacian matrix. Under the assumptions made above (undirected, connected,

non-bipartite graphs), the stationary distribution π of this dynamics is $\pi_i = D_{ii}/2m$. Now we can define the clustered autocovariance matrix of the graph at time t as:

$$R_t = H^T \left[\Pi \exp(-tD^{-1}L) - \pi^T\pi \right] H, \tag{3}$$

with $\Pi = \text{diag}(\pi)$ and the matrix H encoding the partitioning as defined for modularity above. Note that each entry $[R_t]_{ij}$ is the probability for a random walker to start in community i at stationarity and after time t end up in community j minus the probability of two independent walkers to be in i and j at stationarity.

We can then write the *stability of a partition H* at time t as:

$$r(t, H) = \text{trace } R_t. \tag{4}$$

This is a global quality function for a given graph and partition that changes as a function of time. It is easy to see that modularity is a particular case of the linearization of the stability (or of its discrete analog) at time $t = 1$ [2]. The stability $r(t, H)$ can now be optimized in the space of partitions H with any optimization method for graph clustering. In the present work, this has been done with the efficient Louvain method [12]. The effect of time is intuitive: with increasing time the Markov process explores larger regions of the graph, so the Markov time acts as a resolution parameter that enables us to identify community structure at different scales. Communities are identified as subgraphs within which the probability distribution of the process is more contained over a time t than otherwise expected at stationarity. Importantly, stability does not aim to find *the* best partition, but rather tries to reveal relevant clusterings at different scales through the zooming process that occurs naturally through the dynamics. A relevant partition should be both persistent over a comparably long timescale and robust with respect to slight variations in the graph structure and/or the optimization [22, 23]. In order to quantify the robustness of the partitions, we use the variation of information to measure the similarity between partitions [23, 24].

Stability differs from one-step methods in its intrinsic multistep character, as it is based on the exponential of the full adjacency matrix. Furthermore, stability does not introduce any effective assumption towards a block-averaged transition matrix and is thus not biased towards a particular structural model of clique-like communities. The fact that stability uses the diffusion dynamics from each node taking into account walks of all possible lengths, rather then looking at one-step transitions only, has an important bearing on the detection of non clique-like communities. As the Markov time increases, stability is able to find the cohesion within non clique-like communities as the multi-step density between nodes in such a community is increased. The dynamic zooming provided by the Markov process is the key characteristic of the stability framework. Such an approach does not focus on a particular scale but rather provides a means to establish the presence of robust partitions that can appear at any given scale or, indeed, reveal the existence of a multiscale community structure. If no communities are present, e.g. in pure Erdös-Renyi random graphs, stability returns the absence of robust communities at any scale. The effect of the Markov dynamics as a zooming lens, scanning from finer to coarser resolution, provides a means to ameliorate the effect of the resolution and field-of-view limits. We should remark that the Markov time does not impose a dynamic on the network necessarily but can rather be seen as a device to reveal the potential community structure in the network (even if the network does not have an intrinsic dynamics).

Community structure in benchmark graphs and in different applications

A number of benchmark graph models have been proposed for the evaluation of community detection algorithms (for some widely used examples see [7, 15–17]). Nearly all of these benchmarks share a particular notion of community: typically a parameter controls the ratio of the probability for a node to connect within its own community versus the probability to connect to a node outside the community. Some more recent models [25] allow for more heterogeneous degree distributions of the nodes, yet a

community can basically still be described as a *homogeneous* substructure and the community structure may be thought of as a probabilistic realization of a 'clique of cliques':

$$p(i,j) \propto k_i k_j f(C_i, C_j), \tag{5}$$

where k_i, k_j are the degrees of node i and j respectively and f is a function of the associated communities C_i, C_j of i and j alone. Therefore, these random graph models can be thought of as 'mean field' or 'block models', in that intra- and inter-community structure is approximated by average connection properties.

If we think of Markov diffusion processes or flows taking place on these graphs, such networks exhibit fast exploration of all nodes within each community—due to the clique-like structure of the community any node in the community is reached in one step. Although graphs of this type are indeed found in real applications (such as networks constructed from correlation measurements or in some instance of social groupings), a wide spectrum of real networks are not of this form because they do not display an 'all-to-all' connection pattern. An important example of networks that cannot support clique-like connection patterns is that of geographically embedded networks, such as power grids, sensor-networks, river-networks, road- and train-networks and other transport and supply and distribution networks. Similarly, networks containing constraints, in some cases dictated by geometry or by some other cost functions, such as higher-dimensional grid or lattice-like structures originating from physical and biological systems, will not display homogeneous block-like structures in their connectivity patterns. We show some examples of these networks in the Results section below.

This issue is generic in the sense that it applies to graphs that have an intrinsic geometric or constraint-driven structure even if it emerges from an abstract data structure. Examples include graphs with a low doubling dimension [26] which are related to graphs that can be naturally embedded in a low-dimensional Euclidean space. Indeed, it has been conjectured that the Internet autonomous systems network–a prominent example of a network pegged to geographical constraints–has low doubling dimension [27] and can be embedded naturally into a two-dimensional hyperbolic plane [28]. Other networks with a similar structure emerge from the projection of high-dimensional data onto a lower dimension while preserving the local structure of the data [29] or in web graphs with a tendril-like structure in the periphery. In all these networks, the connection probability between nodes is strongly influenced by locality and it tends to be sparser and inhomogeneously distributed within and in between different structural components. Invoking again a diffusion process, these graphs may contain non clique-like communities within which diffusion is slow, such that they cannot be identified with standard structural, one-step community detection methods. The fact that a wide range of real networks are not of a mean-field type and their possible community structure may not be clique-like should be taken into account when comparing community detection algorithms on benchmarks which have been developed with the implicit assumption of clique-like communities.

Results

We now analyze a series of constructive and real-world examples in which communities are significantly different from clique-like structures. For all the results presented below, we run the respective community detection algorithms 100 times with a different random seed and select the best partition found as the community structure. For both modularity and stability, the optimization has been performed with the Louvain algorithm [12]. Infomap has its own implicit optimization method. To assess the robustness of communities obtained by the stability method, we used the variation of information [24] of all solutions found by the optimization.

Constructive examples: Non clique-like communities with low intra-community diffusivity

Before considering real-world networks, we illustrate our ideas with constructive toy examples to exemplify the notion of non clique-like communities. The first example is a 'ring of rings', in which 5 rings of 20 nodes with *strong intra-ring* edges are linked via weak edges to each other (Fig. 1A-C). Given our discussion above, it would be desirable that community detection algorithms should reveal the strongly linked ring communities. Indeed, the strong rings correspond to the notion of community as the equivalent of the connected components when the graph is 'almost disconnected', i.e., in the limit when the weaker edges become non-existent. However, as shown in Figure 1A-B, one-step approaches, such as modularity and Infomap, fail to recover these communities and return 'optimal' clusterings that are overpartitioned with many communities of no individual relevance.

Following our discussion in the Methods section, these results can be understood as a consequence of the locally greedy, one-step characteristics of these algorithms. The Infomap algorithm obtains 18 communities as this algorithm does not have any incentive to create communities that go beyond small ring segments, as such communities have low one-step escape probability while being as locally clique-like as possible. In fact, Infomap fails to recognize the rings even when all edges in the graph (within and between communities) have equal weights. Modularity also optimizes according to one-step transitions and, in the case of the rings, such a short horizon is not enough to identify the rings of length 20 and it returns a partition into 8 communities. This effect becomes more acute when the length of the rings increases, e.g., with 2 rings of 50 nodes modularity finds 10 communities instead of the expected 2.

In contrast, the stability formalism identifies only one persistent and robust partition: the right split into 5 rings (Fig. 1C). Stability detects the communities when the dynamic zooming reaches sufficiently long Markov times. In this case, this occurs at a Markov time greater than 1, thus explaining why modularity is unsuccessful. Indeed, above $t \simeq 2$, a long-lasting plateau of stability corresponding to 5 communities emerges and this partition is robust as given by the vanishing value of the variation of information calculated for 100 optimizations of the stability at each time point.

As explained above, the Markov time at which the structure is detected is indicative of an effective intra-community diffusion distance that needs to be spanned to make the community identifiable. This is akin to a measure of the diameter of the community (specifically in the case of quasi-regular graphs). The fact that both Infomap and modularity are effectively one-step methods means that if the communities they detect are the 'right ones', they should always have low diameter. Conversely, when the communities returned as optimal by Infomap and modularity have large diameters, this can be seen as an indication of an overpartitioning effect by these methods. In that case, both algorithms are operating in a regime for which they were not designed since they will always try to find locally dense or one-step structures and the communities may thus be too fine grained.

This observation can be used to provide a check for the appropriateness of those methods for the particular network analyzed. As a simple proxy, one can monitor the diameter of the detected communities[1]. For instance, in the case of the ring of rings, the corresponding average diameters of the communities found by each method are $\Delta_{\mathrm{Mod}} \approx 14$ and $\Delta_{\mathrm{Map}} \approx 4.56$, indicating a potential overpartitioning. Similar diameters are observed in the real-world examples shown below, as does the fact that the communities detected by Infomap tend to be smaller in diameter than the ones found by modularity, a reflection of Infomap being more greedy towards locally clique-like structure. This feature of Infomap, which makes it immune to the resolution limit, makes Infomap more sensitive to the field-of-view limit, and hence to overpartitioning.

Further insight into the importance of the intra-community distance in non clique-like communities can be gained by considering another constructive example: a ring of 5 small-world (SW) graphs [30] with

---

[1]The diameter is just an easily computable indicator of this effect but other more appropriate measures could be used, specially in the case of non-regular graphs. Here we calculate the average of the maximum of the shortest paths in the subgraphs induced by the partioning.

200 nodes each in which edges inside SWs are five times stronger than edges between SWs (Fig. 1D-F). The SWs are constructed using the CONTEST toolbox [31] as follows: start from a ring with nearest and next-nearest connections and add a random shortcut to each node with probability p. As shown in Figure 1D, both Infomap and modularity suffer from severe overpartitioning: when the SW graphs have few shortcuts, Infomap finds 87 communities and modularity returns 27 communities while stability finds the right split into 5 communities.

As the shortcut probability is increased, the diameter of the SW is reduced while their local structure is basically unaffected going from a 'large world' to a 'small world' (this is the classic finding by Watts and Strogatz [30]). From our viewpoint, this means that the SWs become smaller (i.e., their diameter decreases) and these communities should become easier to detect by one-step methods. Indeed, it is shown in Figure 1E that modularity detects the correct number of communities when their mean diameter falls below 7. On the other hand, Infomap does not detect the SWs as communities in the range shown in Fig. 1E, and only detects the SWs when the diameters fall below a diameter of around 4 (not shown). Again, this highlights the bias of Infomap towards locally clique-like communities with short effective diameters. The same feature that makes Infomap successful in dealing with the resolution limit appears here at the basis of its susceptibility to the field-of-view limit.

Figure 1E also shows that stability consistently detects the right community structure with 5 SWs as the only persistent and robust partition, for all densities of shortcuts, i.e., no matter how small or large their diameter is. According to our dynamic viewpoint, it is expected that as the diameter decreases, the SW communities will be detected at smaller Markov times. This is presented in Figure 1F, where we show that the Markov time at which the SWs become identifiable by stability decays roughly as the inverse of the shortcut probability. Satisfyingly, this correlates well with the dependence of the mixing time in SWs estimated from the spectral gap, i.e., the first non-zero eigenvalue of the Laplacian [32].

Networks with non clique-like communities from diverse real-world applications

As shown above, the key limitation of modularity and Infomap for the analysis of non-clique community structures is their reliance on a single-step (structural) notion of community, which leads to existence of a fixed scale in the algorithm. In contrast, stability zooms across scales and, importantly, the scanning through all scales is an essential feature of the algorithm. By scanning through time and considering all relevant features detected along the way, a multi-scale structure can be revealed, yet without imposing a scale *a priori*. In cases where there is no community structure, as for Erdös-Renyi random graphs, the algorithm does not find structure at any scale. If there is one or more scales, they can be found through the sweeping. Therefore, in stability, the potentially multi-scale structure of the graph is explored through a multi-step process given by the Markov time.

Graphs from image analysis As a first real application we consider an image segmentation problem in which the aim is to analyze an image and find meaningful substructures without *a priori* knowledge or guidance. One approach towards this problem is to create a graph representation of the image, in which each node corresponds to a pixel in the original image and the weight of the edges is computed according to image properties such as distance between the pixels, difference in intensity and/or color, etc. [33]. The resulting graph can then be analyzed for community structure as a means to detect meaningful substructures in the image. Clearly, the graphs thus generated will not have homogeneous, clique-like community structure, since the graph is generated from an image with a two-dimensional structure and it incorporates a distance metric.

In Figure 2 we present the results of the analysis of a sample image [2] of size 102×102 pixel. The

---

[2]The image is freely available in various sizes in png format at http://www.iconarchive.com/show/christmas-icons-by-mohsenfakharian/balloons-icon.html. For the present analysis, a jpg variant of slightly different

graph we analyze is constructed from the grayscale version in Fig. 2A. As expected from our discussion above, Figure 2D-E shows that both Infomap and modularity lead to an over-segmentation of the image into 213 and 37 communities, respectively, with average diameters larger than one: $\Delta_{\mathrm{Map}} \approx 3.67$ and $\Delta_{\mathrm{Mod}} \approx 8.27$, a fact that is indicative of non clique-like communities. On the other hand, the dynamic zooming provided by stability finds a naturally robust partition into 16 communities, which emerges at a Markov time of around 11 (Fig. 2B-C) and corresponds well with the underlying features of the image. The robust partition is indicated both by its persistence in terms of Markov time and by its robustness, as indicated by a minimum in the variation of information between the partitions found by the optimization. For comparison, we also evaluated in Fig. 2F hierarchical Infomap [34], which has been proposed recently and allows to agglomerate hierarchically the communities obtained by the Map equation formalism into different levels. Our analysis shows that only the clusterings obtained at the highest level of hierarchical Infomap provide a perceptual improvement and the hierarchical scheme still leads to communities split into subcomponents of no obvious significance. This example from image segmentation highlights the relevance and significance of a notion of community that deviates from the usual clique-like assumptions, which can only be detected with multi-step community detection algorithms, such as stability optimization.

Protein structure analysis   Proteins are a class of macromolecules with complex three dimensional spatial structure which exhibit a hierarchy of motions that are intimately coupled to their function. Structural analysis that can shed light into their function is a very active area of research. Although it is well known that identifiable motifs appear at different time and length scales, a coherent methodology that can provide an integrated description of the hierarchy of structures from the bottom-up remains elusive. As a second example, we show in Figure 3 the analysis of the protein Adenylate Kinase (AdK), an enzyme which functions by performing an opening and closing global motion at slow timescales. Therefore, AdK has been studied both experimentally and theoretically as a model for the analysis of hierarchical dynamics of proteins and as a benchmark for method development (see, for example [23] and references therein). In this case, the graph is created from structural data, i.e., from the positions of atoms in three-dimensional space and the interactions between them. This results in a graph in which the nodes are atoms and the edges correspond to bonds and chemical constraints (for details see [23]).

The structural and geometric origin of the graph leads to a non clique-like community structure. Again, this causes overpartitioning for both Infomap and modularity, as can be seen in Figure 3A-B: Infomap returns 421 communities with average diameter $\Delta_{\mathrm{Map}} \approx 4.05$ while modularity detects 69 communities with average diameter $\Delta_{\mathrm{Mod}} \approx 10.12$. The large diameters again indicate that both methods are operating in a regime that does not match their intrinsic assumption of what constitutes a good community. Indeed, the partitions obtained by these methods do not reflect the dynamical and structural features that are prominent in AdK [1, 23] and even the highest level in hierarchical Infomap (Fig. 3C) is overpartitioned and does not provide an appropriate coarse graining in this case. On the other hand, the robust partitions detected by stability disclose the multiscale structure of the protein graph, revealing important functional and structural subunits, such as amino acids, secondary structures and conformational substructures as exemplified in Figure 3D-F [1, 23].

Power grid network   Our final example is the analysis of a classical technological network, namely the power grid of Continental Europe [3]. This network is based on data from the Union for the Coordination of Transmission Energy (UCTE) and has been analyzed previously for robustness to targeted attacks [35, 36]. As this network is embedded geographically and is constrained by engineering costs, its structural properties are far from 'all-to-all' connectivity and we expect non clique-like communities.

---

size has been used.

[3]Dataset taken from http://www.termoenergetica.upc.edu/marti/index.htm

The community structures obtained by stability, modularity and Infomap are shown in Figures 4 and 5. The results for Infomap and modularity follow the same pattern as above: Infomap overpartitions into 254 communities with average diameter $\Delta_{\mathrm{Map}} \approx 4.91$ while modularity returns 32 communities with $\Delta_{\mathrm{Mod}} \approx 13.84$ (Fig. 4A-B). Both methods result in a fractured representation of the European power grid. Hierarchical Infomap finds non-meaningful partitions for low hierarchical levels and improved partitions only at its highest level, although still displaying segregated small regions.

Through its intrinsic dynamic sweeping, stability reveals a multiscale structure with meaningful sub-regions of different sizes at various Markov times. The Markov times highlighted in Figure 5 have been selected according to the relative decrease in variation of information, which indicates a more robust partitioning. Interestingly, the communities found by stability at different times appear to be related to historical and commercial features of the power grid network. For larger times, the coarse communities correspond well with big historical monopolies, basically identified with nations. Germany provides an exception to this scheme since the German power grid is split between four large companies (with one covering the eastern part of Germany)[4]. This is reflected faithfully by the community structure detected by stability (Fig. 5D). For shorter times (Fig. 5B-C), we get communities on a sub-national scale that also correspond to regional operators, e.g., France gets split up into several communities which overlap well with the regional organization of the French power grid [5]. Similar effects are observed in Spain, Italy and Switzerland. In Figure 6, we show a representation of the communities found as the dynamic zooming lens of the Markov process is applied to this network. The detailed analysis of these observations will be the object of future work.

The dynamic zooming provided by the Markov time is not only the key ingredient to detect communities without a restriction of scales but also the robustness (or lack of robustness) of specific communities is revealing of characteristics of those communities. In the case of the grid, interesting effects can be found when looking at the way communities coarsen as the Markov time gets larger. For instance, the communities in Switzerland appear to flip between different adjacent communities as the Markov time increases. This lack of robustness indicates a strong shared interconnectedness of the Swiss communities with its neighboring groups. This aligns well with the known fact that the Swiss power-grid is an important mediating hub in the center of Europe with 11% of Europe's electricity flowing through Switzerland [6].

Discussion

The examples considered above illustrate how one-step methods, which are tuned to detect clique-like or low diameter communities, are prone to displaying a form of overpartitioning for networks with non clique-like community structure. This is the result of the existence of a field-of-view limit for structural, one-step algorithms, such as Infomap and modularity. The field-of-view limit establishes an upper limit on the size of the communities that can be detected, as given by their effective intra-community distance. Therefore, non clique-like communities present a challenge because they have an inherent scale that might fall beyond the field-of-view of standard one-step methods. When the intra-community effective distance is larger than the field-of-view, this leads to overpartitioning. The field-of-view appears on the opposite scale of the well-known resolution limit [18], which indicates a lower limit size below which modularity cannot detect communities.

Indeed, Infomap can be seen as having being designed to resolve absolutely the fine scale (hence no resolution limit) by optimizing for locally clique-like substructures. This feature, which makes it extremely successful in the analysis of networks with high density of connections, makes it 'myopic' with respect to larger non clique-like structures. Therefore, Infomap suffers from a large field-of-view

---

[4]See http://de.wikipedia.org/wiki/Stromnetz#Netzbetreiber

[5]For a map of the French regional electrical companies see http://www.rte-france.com/fr/nous-connaitre/qui-sommes-nous/organisation-et-gouvernance/le-siege-et-les-unites-regionales

[6]See https://www.swissgrid.ch/swissgrid/en/home/europa/european_overview/ch_in_europe.html

limit, even more acutely than modularity, as shown in the examples above. Interestingly, the fact that Infomap can also be seen as a one-step method from a dynamical perspective, as discussed above, has allowed us to propose a multi-step correction of the Map formalism that may provide a remedy for this problem [20]. Modularity has an intrinsic scale dependent on the overall size (weight) of the graph which allows it to resolve communities in a range between the resolution and field-of-view limits. Depending on the particular graph, this scale may be well matched to the community structure present (in which case, modularity works) or may be too large (underpartitioning, affected by the resolution limit) or too small (overpartitioning, affected by the field-of-view limit). Stability on the other hand applies a dynamic zooming across all scales following the Markov time evolution and thus does not impose *a priori* a specific scale to detect the community structure. It is important to remark that the dynamic zooming is a key aspect of the analysis: only by scanning through Markov time one can reveal whether a specific (time-)scale is meaningful for the problem at hand or whether it merely corresponds to an artificial, non-robust partition.

Although the standard assumption of communities as probabilistic clique-like groups is well motivated and relevant for a number of important complex systems, our examples highlight the fact that this view of community is not always representative of the structures found in networks of current interest. Segments in an image, functional subunits of a protein or geographic entities in the power network are modular substructures of networks that are only adequately identified when multi-step community detection methods, such as stability optimization, are employed. Furthermore, multi-step methods such as stability can also be used for the analysis of clique-like community structures, as shown in Ref. [37] with benchmark models (see Fig. 2 therein and note that the Reichardt-Bornholdt Potts method [9, 10] has been shown [1, 2] to correspond to a linearization of our partition stability formalism). As recent work suggests [38], however, it may not be possible to resolve highly inhomogeneous community structure at a single (fixed) time, and further research needs to be pursued to develop local methods that do not infer a set of scales from the global graph.

Arguably, communities that are non clique-like are in some instances most relevant, in the sense that they correspond to subsystems in which all parts are related but not necessarily in a direct manner. This is a specific network viewpoint for data analysis, as opposed to a generic set of relationships between elements based on correlation. In fact, the analysis of networks with clique-like communities, corresponding to subsystems where all nodes interact with each other, may be more appropriately pursued through multivariate statistics, rather than explicit network analysis. This realization also has implications for current efforts to construct networks from data where, in contrast to correlation matrices, locally sparser networks are favored as a means to reveal the underlying systemic connections of the data [6].

Nearly all commonly used benchmark models to date have adopted a clique-like notion of community, a fact that needs to be taken into account when considering comparative tests performed on these benchmarks, which are likely to be favorable towards methods that have been designed for the detection of a clique-like notion of community [19]. Although these benchmark graphs reflect adequately the community structure of many important networks and datasets, we have shown that many real-world graphs are likely to have non-clique community substructures. This observation hints at the need for a broader set of benchmarks (e.g., random geometric graphs) that go beyond a block-averaged structure and can reflect the specific properties of networks from different application areas.

No community detection method will be universally optimal for the analysis of all networks and, as pointed above, trade-offs between specific and generic features are to be expected. Our findings highlight the critical importance of a detailed assessment of the assumed definition of community and of its appropriateness to the features of the network to be analyzed. Our work also reinforces the need for a careful consideration of the scales that might be implicit in community detection algorithms that can lead to the identification of non-relevant structures as communities. Our use of the multi-step stability framework proposes to circumvent the implicit assumption of a scale by considering community detection as a zooming process through all scales in a systematic manner, as given by a Markov diffusion on the

graph. Alternatively, one can understand the stability formalism as finding the particular time scale at which modules in the graph will be seen as clique-like communities after the process diffuses on the graph.

However, and specifically in the case of non-regular graphs, there remain issues inherent to the fact that all the methods considered in this paper use global definitions of community. As stated above, there is a need for further research towards local methods for community detection that avoid the restrictions imposed by the global graph scale and potentially allow for a soft, overlapping partitioning of the network. More generally, although a huge set of community detection algorithms has been presented to date, it is still an open problem to establish connections between the underlying themes (conceptual and algorithmic) that these methods implement and the applications for which they might be especially suitable. This will be the object of future work.

Acknowledgments

References

1. Delvenne JC, Yaliraki SN, Barahona M (2010) Stability of graph communities across time scales. Proceedings of the National Academy of Sciences 107: 12755-12760.

2. Lambiotte R, Delvenne JC, Barahona M (2008) Laplacian Dynamics and Multiscale Modular Structure in Networks. URL http://arxiv.org/abs/0812.1770. Arxiv preprint.

3. Fortunato S (2010) Community detection in graphs. Physics Reports 486: 75 - 174.

4. Castellano C, Fortunato S, Loreto V (2009) Statistical physics of social dynamics. Rev Mod Phys 81: 591–646.

5. Rives AW, Galitski T (2003) Modular organization of cellular networks. Proceedings of the National Academy of Sciences of the United States of America 100: 1128-1133.

6. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. Nature Reviews Neuroscience 10: 186–198.

7. Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. Phys Rev E 69: 026113.

8. Newman MEJ (2006) Finding community structure in networks using the eigenvectors of matrices. Phys Rev E 74: 036104.

9. Reichardt J, Bornholdt S (2004) Detecting Fuzzy Community Structures in Complex Networks with a Potts Model. Phys Rev Lett 93: 218701.

10. Reichardt J, Bornholdt S (2006) Statistical mechanics of community detection. Phys Rev E 74: 016110.

11. Ronhovde P, Nussinov Z (2009) Multiresolution community detection for megascale networks by information-based replica correlations. Phys Rev E 80: 016109.

12. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 2008: P10008.

13. Rosvall M, Bergstrom CT (2007) An information-theoretic framework for resolving community structure in complex networks. Proceedings of the National Academy of Sciences 104: 7327-7331.

14. Rosvall M, Bergstrom CT (2008) Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences 105: 1118-1123.

15. Danon L, Díaz-Guilera A, Arenas A (2006) The effect of size heterogeneity on community identification in complex networks. Journal of Statistical Mechanics: Theory and Experiment 2006: P11010.

16. Lancichinetti A, Fortunato S, Radicchi F (2008) Benchmark graphs for testing community detection algorithms. Phys Rev E 78: 046110.

17. Lancichinetti A, Fortunato S (2009) Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. Phys Rev E 80: 016118.

18. Fortunato S, Barthélemy M (2007) Resolution limit in community detection. Proceedings of the National Academy of Sciences 104: 36-41.

19. Lancichinetti A, Fortunato S (2009) Community detection algorithms: A comparative analysis. Phys Rev E 80: 056117.

20. Schaub MT, Lambiotte R, Barahona M (2011) Coding of Markov dynamics for multiscale community detection in complex networks. Arxiv preprint arXiv:1109.6642v1.

21. Newman MEJ (2006) Modularity and community structure in networks. Proceedings of the National Academy of Sciences 103: 8577-8582.

22. Lambiotte R (2010) Multi-scale modularity in complex networks. In: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2010 Proceedings of the 8th International Symposium on. IEEE, pp. 546–553. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5520347.

23. Delmotte A, Tate EW, Yaliraki SN, Barahona M (2011) Protein multi-scale organization through graph partitioning and robustness analysis: application to the myosin:myosin light chain interaction. Physical Biology 8: 055010.

24. Meila M (2007) Comparing clusterings–an information based distance. Journal of Multivariate Analysis 98: 873 - 895.

25. Karrer B, Newman MEJ (2011) Stochastic blockmodels and community structure in networks. Phys Rev E 83: 016107.

26. Gupta A, Krauthgamer R, Lee J (2003) Bounded geometries, fractals, and low-distortion embeddings. In: Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on. pp. 534 - 543. doi:10.1109/SFCS.2003.1238226.

27. Fraigniaud P, Lebhar E, Viennot L (2008) The Inframetric Model for the Internet. In: 27th IEEE International Conference on Computer Communications (INFOCOM). IEEE, pp. 1085-1093.

28. Boguna M, Papadopoulos F, Krioukov D (2010) Sustaining the Internet with Hyperbolic Mapping. Nature Communications 1: 62.

29. Tenenbaum JB, Silva Vd, Langford JC (2000) A Global Geometric Framework for Nonlinear Dimensionality Reduction. Science 290: 2319-2323.

30. Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393: 440–442.

31. Taylor A, Higham DJ (2009) CONTEST: A Controllable Test Matrix Toolbox for MATLAB. ACM Trans Math Softw 35: 26:1–26:17.

32. Barahona M, Pecora LM (2002) Synchronization in Small-World Systems. Phys Rev Lett 89: 054101.

33. Browet A, Absil P, Van Dooren P (2011) Community Detection for Hierarchical Image Segmentation. In: Aggarwal J, Barneva R, Brimkov V, Koroutchev K, Korutcheva E, editors, Combinatorial Image Analysis, Springer Berlin / Heidelberg, volume 6636 of *Lecture Notes in Computer Science*. pp. 358-371. URL http://dx.doi.org/10.1007/978-3-642-21073-0_32.

34. Rosvall M, Bergstrom CT (2011) Multilevel Compression of Random Walks on Networks Reveals Hierarchical Organization in Large Integrated Systems. PLoS ONE 6: e18209.

35. Rosas-Casals M, Valverde S, Solé RV (2007) Topological Vulnerability of the European Power Grid under Errors and Attacks. I J Bifurcation and Chaos 17: 2465-2475.

36. Solé RV, Rosas-Casals M, Corominas-Murtra B, Valverde S (2008) Robustness of the European power grids under intentional attack. Phys Rev E 77: 026102.

37. Traag VA, Van Dooren P, Nesterov Y (2011) Narrow scope for resolution-limit-free community detection. Phys Rev E 84: 016114.

38. Lancichinetti A, Fortunato S (2011) Limits of modularity maximization in community detection. Phys Rev E 84: 066122.
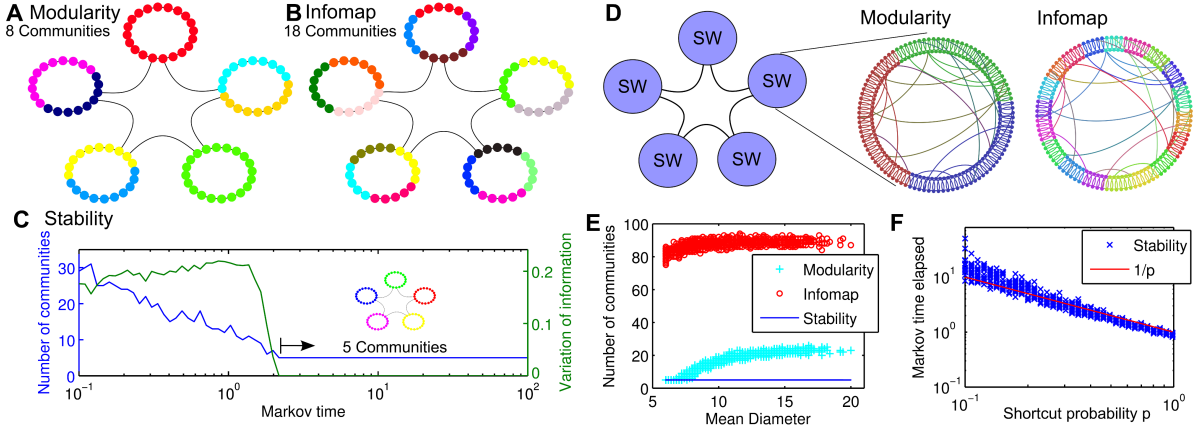
Figure Legends

Figure 1. Constructive networks with non clique-like community structure. **A-C** Ring of rings: 5 rings of 20 nodes in a ring configuration. The edges within each ring are 5 times stronger than between rings. Optimal communities according to: **A** modularity (8 communities found) and **B** Infomap (18 communities found). **C** Analysis with stability: number of communities found (blue) and average variation of information (VI) of the partitions found (green) as a function of Markov time. The average VI is obtained from 100 runs of the Louvain algorithm. Starting at $t \approx 1.96$ the correct partitioning into 5 communities is detected as a persistent, robust partition with vanishing VI. No other stable partition is detected at any other Markov time (as shown by the high values of VI). **D-F** Ring of small worlds: 5 SWs of 200 nodes in a ring configuration. The edges inside each SW are 5 times stronger than those between different SWs. The SW property, and thus the diameter of the SWs, is varied by varying the shortcut probability $p \in [0.1, 1]$ (see text for details). **D** Examples of partitions obtained by modularity and Infomap for $p = 0.1$. **E** Number of communities detected as a function of the measured mean diameter of the small world subgraphs for Infomap, modularity and stability: Infomap never detects the SWs; modularity detects the SWs only when they have short enough diameter; stability always detects the SWs (at different Markov times) as the only stable partition. **F** Markov time elapsed until the SWs are detected by stability as a function of the shortcut probability.
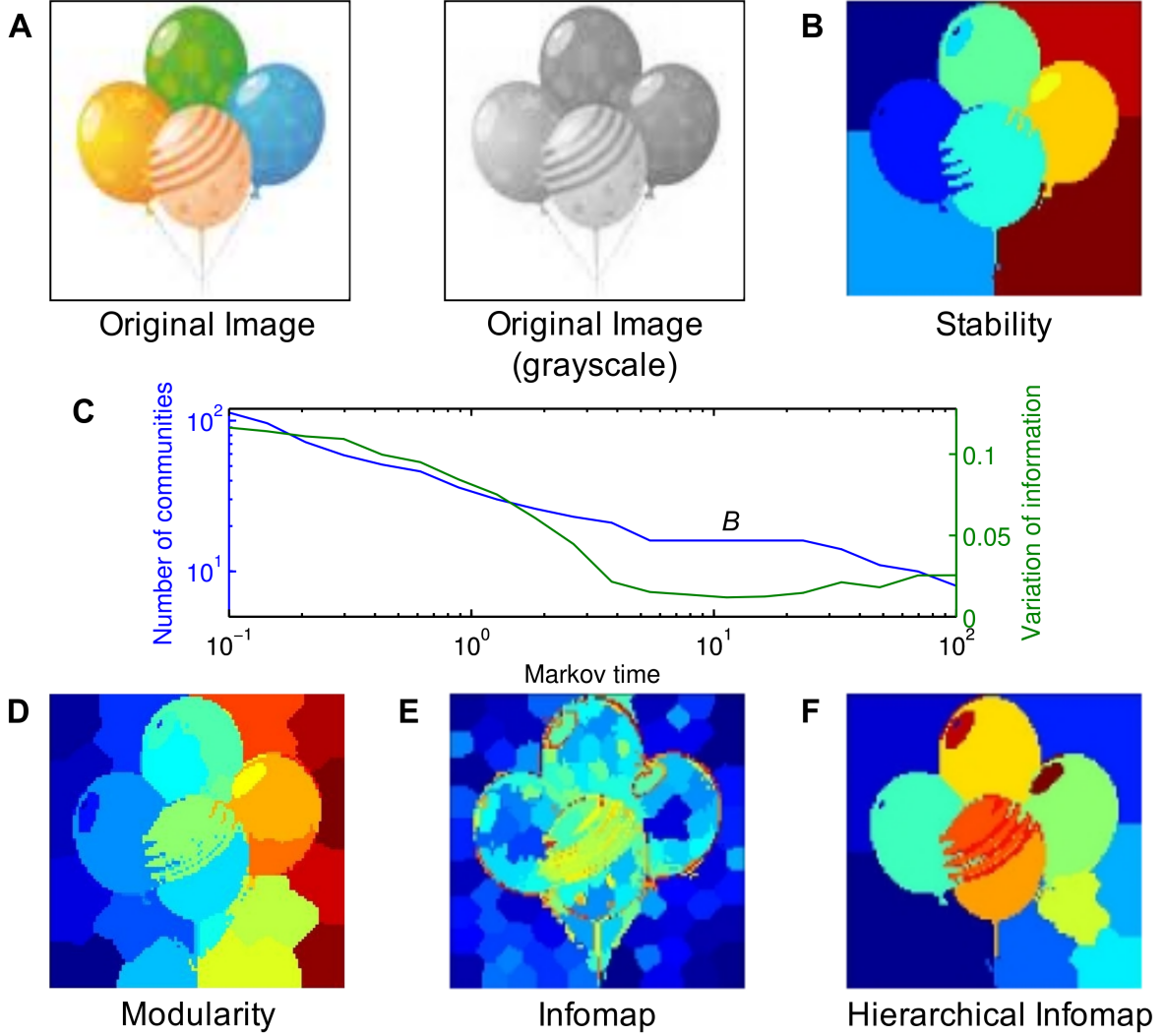
Figure 2. Image segmentation via community detection. **A** Original image in color and the grayscale version used in the analysis. **B** Image segmentation (false color plot) found with stability (16 communities). **C** Stability analysis of the image graph. The partition in **B** corresponds to the plateau at Markov time $t = 11.3$, where a minimum of the VI also occurs. **D-F** Image segmentation (false color plot) obtained from different community detection methods: **D** modularity (37 communities), **E** Infomap (213 communities), **F** hierarchical Infomap (15 communities; highest hierarchical level).
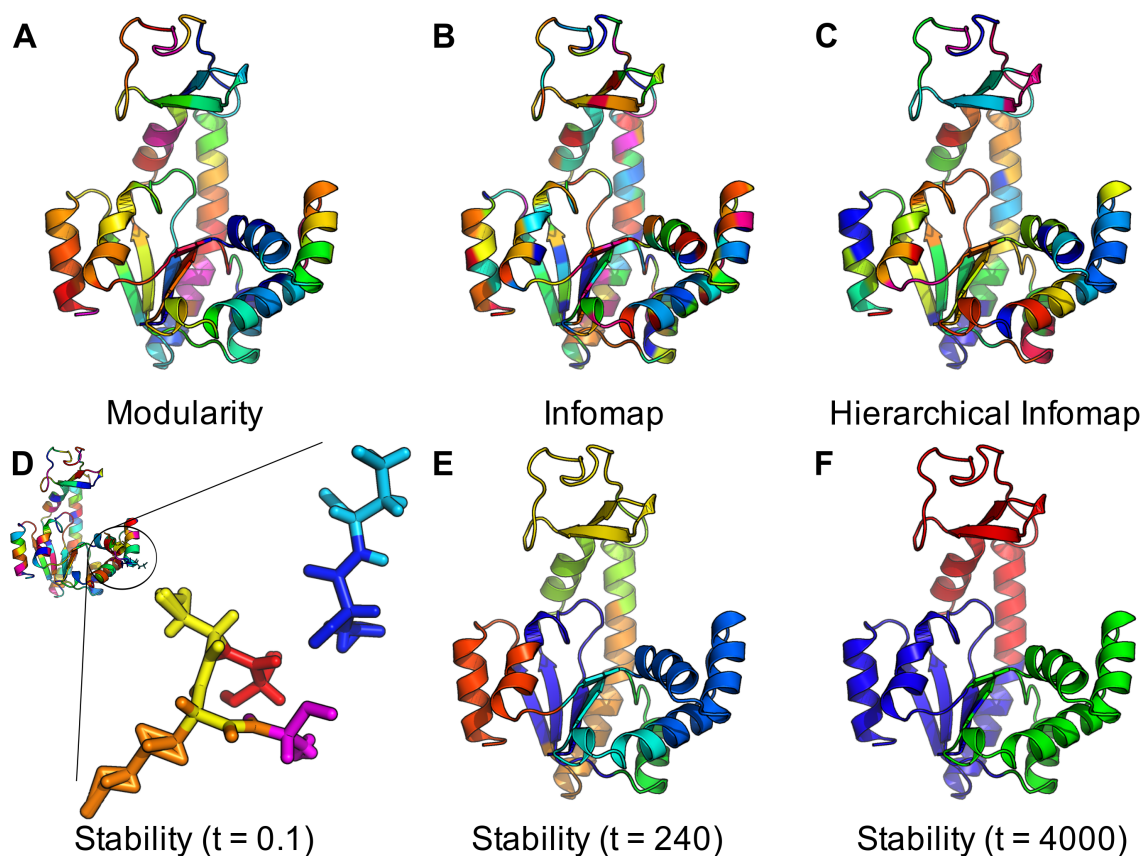
Figure 3. Analysis of the structural graph of the protein Adenylate Kinase (AdK). A-E Visualization of the communities found by the different algorithms (adjacent regions in the same color correspond to communities): **A** Modularity (69 communities); **B** Infomap (421 communities); **C** hierarchical Infomap (58 communities). **D-F** Some of the robust communities found by stability at different Markov times: **D** $t = 0.1$ (206 communities), the communities capture the amino acids of the protein (214 amino acids); **E** $t = 240$ (8 communities), the communities correspond approximately to secondary structure of the protein (c.f. [23]); **F** $t = 4000$ (3 communities), the communities correspond to the functional domains of the protein that operate at slow timescales. Note that stability finds meaningful substructures also for other times not shown [1, 23].
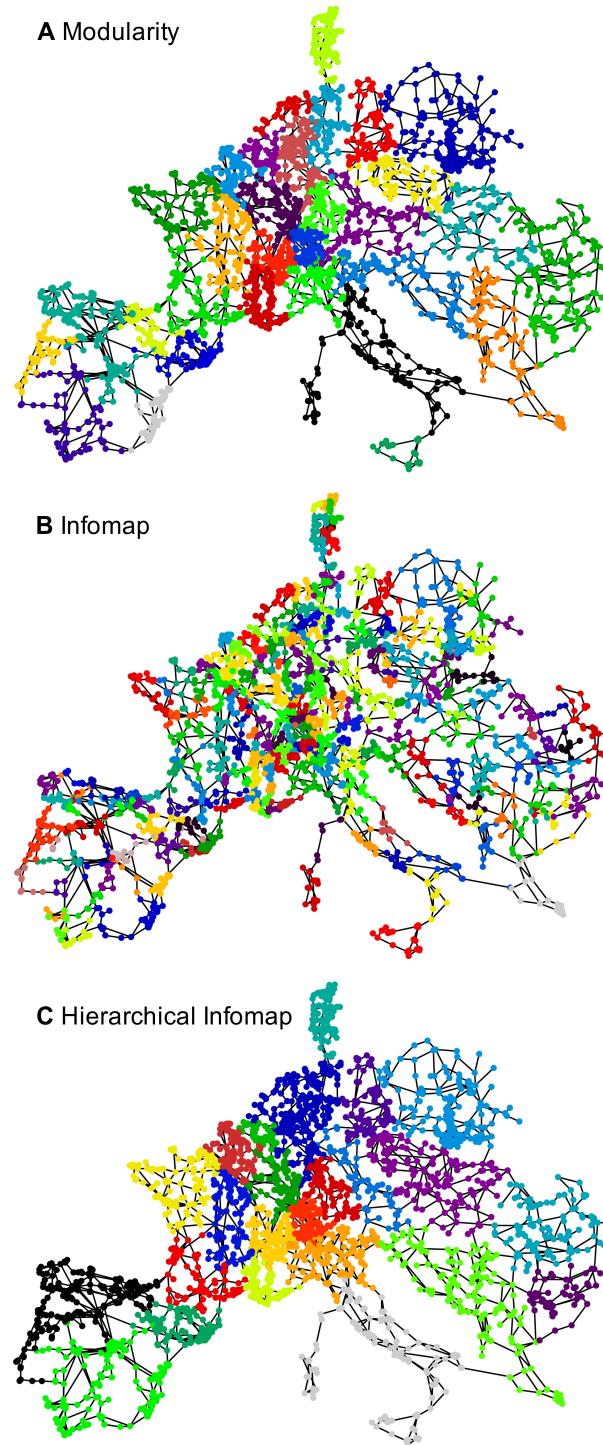
Figure 4. Community structure analysis of the European power grid with different one-step methods: **A** modularity (32 communities); **B** Infomap (254 communities); **C** hierarchical Infomap (24 communities; top level of hierarchy).
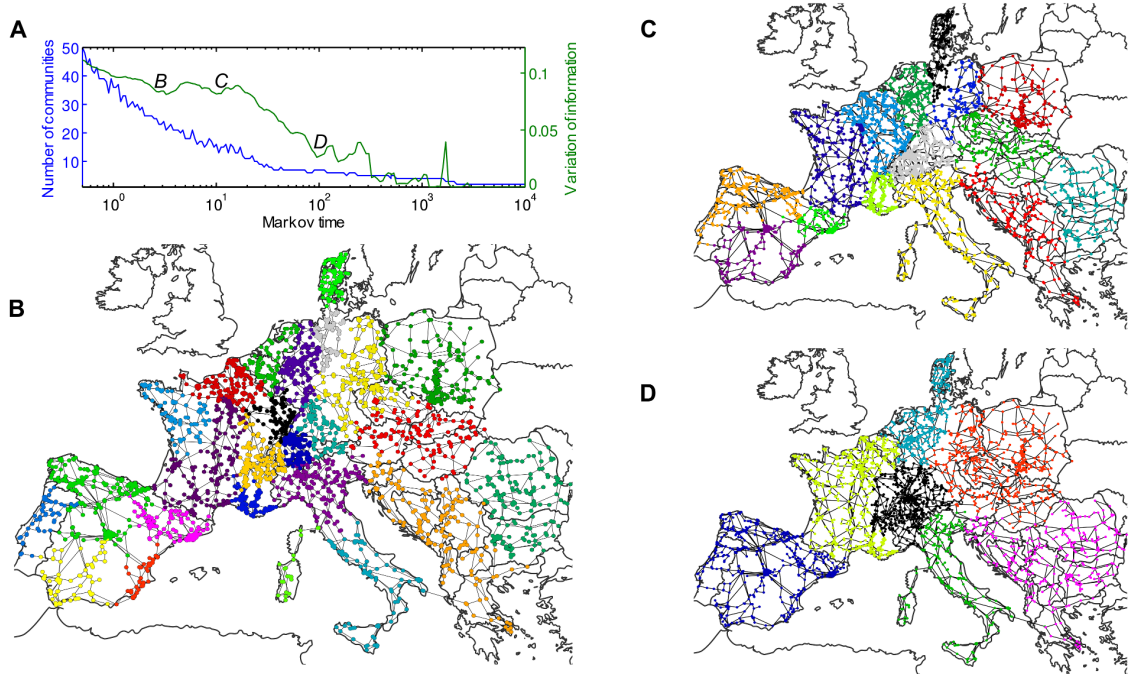
Figure 5. Community structure analysis of the European power grid with stability. **A**
Stability analysis of the community structure of the power grid graph: number of communities found
(blue) and average VI (green) vs Markov time. In this case, 1000 initializations of the Louvain algorithm
have been used to find the best community and compute the variation of information. **B-D** Stability
finds robust partitions for different Markov times that seem to be related to known structure in the power
grid. The partitions shown correspond to Markov times: **B** $t = 2.63$ (25 communities), **C** $t = 11.76$ (15
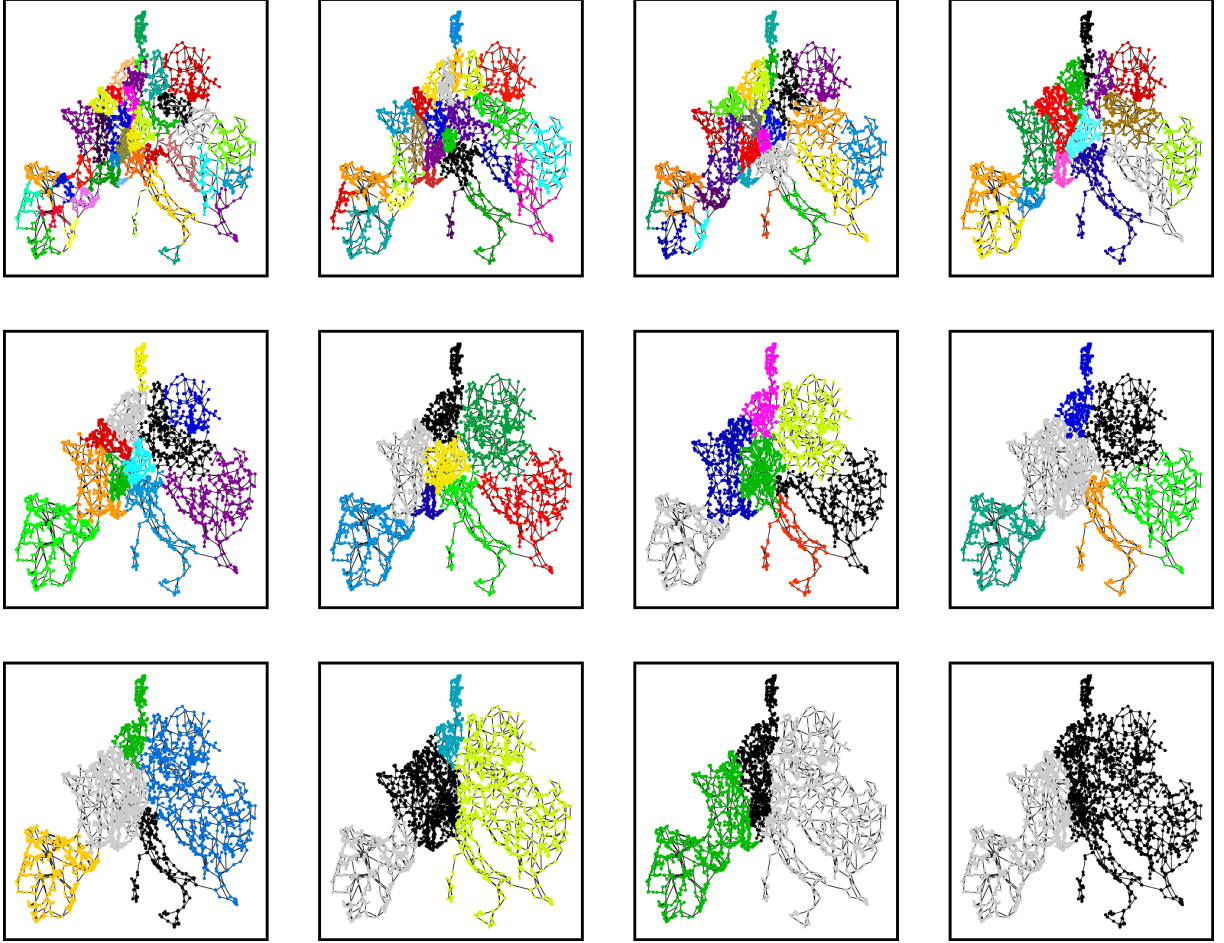communities) and **D** $t = 94.79$ (7 communities).

Figure 6. **Multiscale community structure of the European power grid with stability**. The illustrative partitions shown correspond to Markov times $t = \{0.81, 2.12, 2.62, 11.75, 20.08, 34.29, 94.79, 173.49, 389.77, 1176.77, 1919.14, 10000\}$ (from left to right, top to bottom) and have been selected based on their relative robustness. The dynamic zooming provided by the Markov time provides a progressively coarser representation of the network that captures geopolitical and commercial features of the grid.